

The Impact of TV Ads on the Individual User's Online Purchasing Behavior

Completed Research Paper

Martin Stange

Leuphana University Lüneburg

Lüneburg, Germany

martin.stange@uni.leuphana.de

Abstract

The importance of a well-balanced cross-channel marketing strategy has increased over the past few years. The synergies caused by the interdependencies of different online channels, such as e-mail advertising, search engine and banner advertising, have also drawn the attention of many researchers. However, relatively little is known about the impact of offline marketing, such as TV advertising, on online user behavior. In this article, a model commonly used in clickstream analysis is extended by adding several TV advertising variables. Based on this model, a hierarchical Bayesian logistic model is developed to estimate the cross-channel effects of both offline and online channel contacts. By applying this model to a case study, it is shown which online channels are most supported by television ads. The findings of this paper have managerial implications for practitioners in the field, in particular because of the increasing use of a so-called "second screen" while watching TV.

Keywords: Decision aids, User behavior, Electronic commerce, Bayesian network

Introduction

Data sets produced in online marketing make it possible to analyze advertising effects on the individual user's level (Bucklin and Sismeiro 2009). Based on such analyses, marketing budgets can be attributed to individual online channels, such as affiliate marketing, search engine advertising (SEA) or e-mail advertising (Shao and Li 2011). In practice, however, rather simple heuristics are typically used to determine the marketing success of the individual advertising channels (Dalessandro and Perlich 2012). Thus, such simple methods do not necessarily measure the actual effects of the channels. Consequently, analytical solutions for advertisers are becoming increasingly popular, in particular in the emerging field of real-time advertising, where the probability for a "conversion" or "click" can be used to determine the size of the bid for a given advertising slot (Stange and Funk 2014). To estimate the probability for a conversion or a click, many publications use clickstream data, which is usually then transformed into user journeys (Bucklin and Sismeiro 2009; Chatterjee et al. 2003; Nottorf 2014). The results of user journey analyses can be used to apportion the marketing budget across the channels and to predict the user behavior of future user contacts.

However, when it comes to the interaction of offline marketing campaigns and online user behavior, very few analytical approaches are available. The problem is clear: It is almost impossible to track whether a user visiting a website has been watching a television ad or listening to radio ads in the past hour, day or week. This makes it difficult to apply analytical methods and to attribute the conversions that are achieved online to offline ads.

Nowadays, an increasing number of customers use more than one media device at a time (Courtois and D'heer 2012). For instance, many users surf on the Internet using their tablet or smart phone while watching TV. This use of a so-called "second screen" presents new opportunities for marketers, whose aim it is to reach their customers in a complementary way to maximize marketing effects. Thus, a detailed analysis of a combined online and offline marketing strategy is extremely desirable for many advertisers.

The central research question of this study is how television advertising affects the online purchasing behavior of customers. More specifically, the study investigates how the time dependency of television advertising effects can be modeled and how these time dependent television advertising effects can be included in a commonly used clickstream model to reveal cross-media advertising effects. The study contributes to IS research by providing insights into users' tendency to use different information sources to make decisions in e-commerce contexts. In addition, it encourages IS researchers to focus on this dynamic research field at the intersection of IS and marketing, and for instance, investigate users' multi-device usage or develop strategic decision support systems and real-time bidding agents using probabilistic models (Stange and Funk 2014) such as the one presented in this study. In addition, they can develop new business models for marketing agencies and other service providers in the field (Veit et al. 2014). These models are important for companies, because they often focus only on either online or offline decision support services (Joo et al. 2014).

The modeling approach presented in this paper is structured as follows: First, the clickstream model introduced by Chatterjee et al. (2003) is extended by introducing non-linearity terms to model saturation effects. These changes allow for a more detailed answer to the research questions. Second, a model to estimate the time dependency of the TV advertising effects is developed. Third, a hierarchical Bayesian logistic regression model is developed to estimate the interdependencies of online and offline channels. Applied to a case study, it is shown that these measures decrease the residual deviance of the fit and increase the predictive accuracy in comparison to the original model. The analysis shows different effects for each TV station and advertised product and delivers the interdependencies of the individual online channels and the TV advertising effects.

The remainder of this paper is structured as follows: First, relevant work on clickstream data analysis in the literature is summarized and different approaches for the analysis of offline-online advertising effects are presented. Second, the data set of the case study and the applied modeling process are described. Third, the results of the analyses are presented in detail. Finally, the findings are summarized and their managerial implications are outlined.

Related Work

This article uses the achievements of two research streams. For the case study presented in this paper, models from previous clickstream analyses are used and combined with modeling approaches from studies dealing with the effects of TV advertising.

Clickstream and Cross-Channel Advertising

Clickstream data consists of data records produced when users interact with an advertiser on the Internet (Bucklin and Sismeiro 2003). This kind of contact might be a click on a display ad, a search request, or any activity on the advertiser's website. The complete set of contacts of a user with the marketing channels of an advertiser is referred to as their user journey.

More than a decade ago, Chatterjee et al. (2003) developed a model to predict a user's individual click proneness based on clickstream data. They derived their model from several theoretical considerations about the effect of display advertisement and used several models to explain the dependent variable $Pr(\text{Click} = 1 | \text{UserJourney})$, i.e., the probability for a click given the user's current contact and the previous contacts. They show that a logistic model with heterogeneity terms across users and user sessions best describes user behavior and discuss how the results from the logistic regression can be interpreted as the different effects of marketing activities (Cho 2003; Lee et al. 2012; Nottorf 2014; Richardson et al. 2007). The resulting knowledge about the effect of the individual channels is valuable for distributing the appropriate amount of money across the marketing channels, i.e., for the budget allocation.

The clickstream modeling approach by Chatterjee et al. (2003) is based on counting the channel contacts within and across user sessions. This procedure assumes a linear relationship between the marketing effect and the logit of the click probability. To account for the short-term and long-term effects of the contacts of the customer with the advertiser, such as display views, Chatterjee et al. (2003) introduce two sets of variables containing the total number of contacts in the current session and in earlier sessions. To account for wear-in effects (the customer growing more and more aware of the product) and wear-

out effects (customer awareness becoming saturated and decreasing with the time), they introduce the quadratic form of some channel variables. However, they do not observe significant effects. In this paper, the channel contacts are modeled as non-linear, i.e., instead of counting the channel variables, a function $f(x) = x^\gamma$ for each covariate is introduced and the exponents γ_j are estimated in a separate step before conducting the main hierarchical analysis. It is shown that the predictive accuracy can be increased significantly by adding this step.

Ever since the work of Chatterjee et al. (2003), this topic has drawn the interest of many other researchers, not least because often, in actual practice, rather simplistic heuristics are used to determine the budget for each channel (Anderl et al. 2014; Jordan et al. 2011; Kitts et al. 2010). An important finding from many studies using different methods is that effects of marketing activities are interdependent (Anderl et al. 2014; Chatterjee 2010; Ghose and Yang 2010; Klapdor 2013; Nottorf 2014; Piercy 2012). That is to say, marketing channels cannot be analyzed separately, but have to be seen in the context of other channels. For this reason, analyses should not only focus on individual channel effects but also on cross-channel effects. This has been done by Ghose and Yang (2010) for the search engine advertising channel and the organic search channel, but to the best of the author's knowledge no published study has investigated the complete set of interdependencies across all advertising channels yet.

However, cross-channel marketing is not limited to online channels (Dinner et al. 2014; Kitts et al. 2010; Olbrich and Schultz 2014). Joo et al. (2014) show that offline data from TV advertising spots can be used to predict customers' online search behavior. The more brand-related TV spots are broadcast, the more users tend to search for these brands. However, modeling offline advertising effects on the individual user's level is not as simple as modeling online advertising contacts, since it is almost impossible to determine if a user, in fact, was exposed to the offline advertisement. However, in this study, a new approach to include the offline variables into the user journey is provided. This approach has not been taken in the literature before.

This paper contributes to the field of clickstream analysis in two ways. First, it presents a new way to model and estimate non-linearity parameters and demonstrates how to include TV advertising data in user journeys. Second, it shows how to use a hierarchical Bayesian logistic model to determine cross-channel effects for all marketing channels for a given user journey data set.

Effect of TV Advertising on Online Behavior

In principle, there are two possible approaches to measuring the effectiveness of TV advertising on online user behavior. First, the so-called advertising stock for a product can be calculated as a function of the frequency of the spots and their reaches. This advertising stock can be used as a measure of the long-term (awareness) effects of TV advertising (Lodish et al. 1995). Second, the direct, performance-oriented impact of a spot can be measured by observing the uplift of page impressions and conversions in the few minutes after the broadcasting of the spot (Lewis and Reiley 2013; Zigmond and Stipp 2010). This study focuses on the latter.

As stated above, the main difficulty in measuring the effect of TV advertising on online behavior is the lack of an indication if a user actually was exposed to the advertisement or not. A way to gain this information is described by Kitts et al. (2010) who propose displaying discount codes during TV spots, codes which can be entered online to receive a discount. Conversions that result from this technique can easily be attributed to TV advertising. However, most campaigns have not followed this approach.

The simultaneous use of two devices, e.g., television or a mobile device, has been the subject of increasing interest over the past five years. A recent survey showed that approximately 30% of television consumers use a mobile device at some point during the day (Bolten 2014). This number indicates that the phenomenon of second-screen use is a significant new opportunity for advertisers to optimize their cross-channel budget allocation and to increase profit due to marketing campaigns.

Today, the analysis of television advertising effects on online user behavior is often driven by very simple heuristics. For instance, website traffic is observed in the minutes before and after a spot is broadcast on TV. The uplift in user traffic resulting from the spot, i.e., the difference of the number of page impressions, can then be assigned to the impact of TV advertising. It is readily apparent that this heuristic can only approximate true TV impact, because it ignores any other effects that might cause the uplift at that

moment, effects which cannot be neglected, particularly by bigger online shops. A more sophisticated method is to observe website traffic over longer periods of time, including periods with and without TV advertising broadcasts. The resulting website traffic can be decomposed using a time-series analysis to obtain the periodic effects (i.e., daily and weekly traffic patterns) and the trend effects indicating the long-term effects from TV advertising and the residual effects indicating the short term effects, which can be attributed to TV advertising.

Investigations of TV advertising on online user behavior are not frequent in the literature. However, there are some researchers who analyze the impact of TV ads on an aggregated level. Lewis and Reiley (2013) and Zigmond and Stipp (2010) analyzed the effect of TV ads during the Super Bowl and the opening show of the Olympic Games, respectively. They see a significant increase in terms of in page impressions and conversions for advertised brands in the minutes after the TV spot. The diagram representing the uplift of visits and conversions after the spot has roughly the form of a Γ distribution, which is relatively intuitive: In the first few seconds after the spot, only a few users use their second screen to search for the product, but after a few minutes a maximum is reached. Afterwards, the direct effect of the advertisement decreases. The limitation that a significant uplift can only be observed for spots with a very high reach is clearly evident here, and the application of this approach to smaller reaches is limited.

A more analytical approach is used by Joo et al. (2014), who measured the effect of TV spots on online search behavior. In contrast to the studies discussed in the previous paragraphs, they also accounted for long-term effects using exponential decay to explain the advertising stock. They observed an increase of brand search requests after a TV advertising broadcast. Liaukonyte et al. (2015) measured the impact of TV ads on the purchasing behavior of customers on an aggregated level. They find that the TV advertising effects vary from spot to spot. However, in contrast to their analysis, this paper focuses on the conversion probability on the level of the individual users and also includes other marketing channels such as display and e-mail advertising. This appears to be the first study that analyzes the cross-channel effects of television advertising on the level of the individual user.

Data Description and User Journey Modeling

In this paper, a data set from a German retailer who operates both online and offline stores is used. The available data set contains approximately 60 million records. Each record included in the raw data represents an instance of a user contacting the website and most of these instances (approximately 90%) are generated by user activity on the retailer's website. These activities include, but are not limited to, viewing product pages, or shipping or payment information, or purchasing of products. The advertising channel that customers used to reach the retailer's website is identified in the first record generated for each session in the data set (10% of the complete data set). This channel may be a display banner, a link in an e-mail, a search engine ad, a social media ad, or an affiliate ad. Additional advertising channels are organic search requests, price engine search requests, direct type-ins of the shop URL in the browser, and direct referrals from another website. The case study period includes the 24 days before Christmas Eve, 2013. This period is most likely characterized by many instances of spontaneous gift shopping. The sample was chosen, because a relatively high spillover effect from TV ads to the online user behavior is expected during that time. The clickstream data and the TV advertising data were collected by two different service providers and provided by the retailer. More information about the experimental setting cannot be given due to non-disclosure agreements.

Using this data and closely following Chatterjee et al. (2003), user journeys are built. These consist of three different types of channel variables and additional control variables. First, the intercept terms $I_{(\cdot)} \in \{0, 1\}$ indicate the type of the current contact. For instance $I_{SM} = 1$ indicates, that the current contact is caused by a click on a social media ad. Second, the session variables $J_{(\cdot)} \in \mathbb{N}_0$ indicate the previous number of channel contacts within the current user session. User activity after one hour of no activity defines a new user session (Chatterjee et al. 2003). For example $J_{SEA} = 2$ indicates that the user has clicked twice on a search engine ad within the current session. Third, the variables $K_{(\cdot)} \in \mathbb{N}_0$ indicate the number of contacts with certain channels in previous user sessions. For instance, $K_{OS} = 5$ indicates that a user has searched five times for the product or brand in previous user sessions. The additional control variables are the number of previous conversions within the current session (CWS) and in previous sessions (CAS), the session number (SN), the inter-session time (IST), the time of the day terms (t, t^2, t^3 and t^4) and an indicator whether the contact has occurred on a weekend or not ($WE \in$

$\{0, 1\}$). The variance of the inter-session time is reduced by calculating the logarithm of the time between two sessions (Chatterjee et al. 2003). The dependent variable for each contact indicates a conversion ($Conv = 1$) or no conversion ($Conv = 0$). An example of a user journey is shown in Table 1 (Stange and Funk 2015).

Table 1. User journey example.											
Contact No.	I_0	I_{SEA}	I_D	J_{SEA}	J_D	K_{SEA}	K_D	CAS	IST	SN	Conv
1	1	1	0	0	0	0	0	0	0 h	1	0
2	1	0	1	1	0	0	0	0	0 h	1	0
3	1	0	1	0	0	1	1	0	6 h	2	0
4	1	1	0	0	1	1	1	0	6 h	2	0
5	1	1	0	1	1	1	1	0	6 h	2	1
6	1	0	1	0	0	3	2	1	2 h	3	0

The set of variables for the j^{th} contact of the i^{th} user is given by the j^{th} entry of the design matrix X_i for user i as stated in Equation 1. An overview of the indices and variable types is given in Table 2.

$$X_{ij} = \{I_0, I_{OS}, I_{TI}, I_A, I_D, I_{SEA}, I_{SM}, I_{EM}, I_{PS}, J_{TI}, J_A, J_D, J_{SEA}, J_{SM}, J_{EM}, J_{PS}, J_R, K_{OS}, K_{TI}, K_A, K_D, K_{SEA}, K_{SM}, K_{EM}, K_{PS}, K_R, SN, IST, CWS, CAS, t, t^2, t^3, t^4, WE\}_{ij} \quad (1)$$

Table 2. Overview of variables and indices.			
Index	Meaning	Variable	Meaning
A	Affiliate marketing	CWS	Number of conversions in current session
D	Display advertising	CAS	Number of conversions across sessions
EM	E-mail advertising	IST	Time between two sessions
OS	Organic search	$OffB_L$	Offline Brand Spot Nr. L
TI	Direct type-in	$OffP_K$	Offline Product Spot Nr. K
PS	Price search	OnB_M	Online Brand Spot Nr. M
R	Direct referral	SN	Session Number
SEA, SA	Search-engine advertising	t	Hour of the Day
SM	Social media advertising	TV_S	TV Station S
-	-	WE	Weekend (Yes/No)

To extend the clickstream model of Chatterjee et al. (2003), the impact of the TV advertisement is modeled by additional variables TV_S^K that hold the time (in minutes) since the last spot of a certain type K was broadcast on a given TV station S . Only spots that have been run in the ten hours previous to the contact are included, since short-term effects from spots ran even earlier are unlikely (Zigmond and Stipp 2010). Individual TV stations are distinguished, as are individual spots (Liaukonyte et al. 2015). In addition, TV ads are categorized into brand-related spots (B) and product-related spots (P) as well as offline-related spots (Off) and online-related spots (On). Spot placement or the reaches of the spot cannot be modeled, because this information is unavailable. This is discussed in detail in the limitation section. The design matrix for the TV effect Y_i holds the TV variables for each contact j of user i as stated in Equation 2. As an example, the variable TV_6^{OnB1} holds the time difference in minutes between the time of the contact and the time of the previous broadcast of the online-related brand-related spot #1 on station #6.

$$Y_{ij} = \{TV_2^{OnB_1}, TV_4^{OnB_1}, TV_6^{OnB_1}, TV_{11}^{OnB_1}, TV_5^{OffP_1}, TV_7^{OffP_1}, TV_8^{OffP_1}, TV_1^{OffB_1}, TV_2^{OffB_1}, TV_3^{OffB_1}, TV_4^{OffB_1}, TV_5^{OffB_1}, TV_6^{OffB_1}, TV_9^{OffB_1}, TV_{10}^{OffB_1}, TV_{11}^{OffB_1}, TV_1^{OnB_2}, TV_2^{OnB_2}, TV_3^{OnB_2}, TV_5^{OnB_2}, TV_{11}^{OnB_2}\}_{ij} \quad (2)$$

Table 3 presents the descriptive statistics of the transformed user journey data sample. These numbers suggest the relative frequency of the channel contacts. All user journeys including only one onsite contact and those longer than 50 contacts were removed from the data set to exclude click robots or other Internet fraud. The statistic is based on a random sample of 3,314,920 advertising channel contacts by 898,796 users. The total number of conversions in this set is 138,437.

Table 3. Descriptive statistics of the complete sample of user journeys.							
Parameter	μ	σ	max.	Parameter	μ	σ	max.
I_{OS}	0.464	0.499	1.000	IST	3.797	3.966	11.000
I_{TI}	0.219	0.414	1.000	CWS	0.005	0.073	8.000
I_A	0.080	0.271	1.000	CAS	0.094	0.367	23.000
I_D	0.037	0.189	1.000	WE	0.293	0.455	1.000
I_{SEA}	0.068	0.252	1.000	$TV_2^{OnB_1}$	35.792	86.364	599.000
I_{SM}	0.004	0.060	1.000	$TV_4^{OnB_1}$	42.331	95.597	599.000
I_{EM}	0.080	0.272	1.000	$TV_6^{OnB_1}$	41.818	86.259	599.000
I_{PS}	0.008	0.087	1.000	$TV_{11}^{OnB_1}$	31.971	71.381	599.000
J_{TI}	0.047	0.282	30.000	$TV_5^{OffP_1}$	85.018	125.020	599.000
J_A	0.111	0.772	46.000	$TV_7^{OffP_1}$	61.819	133.435	599.000
J_D	0.018	0.213	22.000	$TV_8^{OffP_1}$	64.053	140.807	599.000
J_{SEA}	0.052	0.398	31.000	$TV_1^{OffB_1}$	101.830	147.150	599.000
J_{SM}	0.003	0.111	29.000	$TV_2^{OffB_1}$	88.921	138.856	599.000
J_{EM}	0.049	0.380	27.000	$TV_3^{OffB_1}$	90.356	142.837	599.000
J_{PS}	0.007	0.151	20.000	$TV_4^{OffB_1}$	56.931	130.628	599.000
J_R	0.011	0.150	21.000	$TV_5^{OffB_1}$	79.487	135.967	599.000
K_{OS}	0.822	2.464	49.000	$TV_6^{OffB_1}$	56.879	128.091	599.000
K_{TI}	0.982	3.184	49.000	$TV_9^{OffB_1}$	49.614	116.594	599.000
K_A	0.225	1.327	47.000	$TV_{10}^{OffB_1}$	59.392	138.278	599.000
K_D	0.097	0.646	37.000	$TV_{11}^{OffB_1}$	95.170	136.942	599.000
K_{SEA}	0.171	1.238	47.000	$TV_1^{OnB_2}$	63.755	129.567	599.000
K_{SM}	0.018	0.320	41.000	$TV_2^{OnB_2}$	80.297	121.517	599.000
K_{EM}	0.265	1.408	48.000	$TV_3^{OnB_2}$	42.329	101.559	599.000
K_{PS}	0.011	0.192	24.000	$TV_5^{OnB_2}$	61.446	109.081	599.000
K_R	0.099	0.790	44.000	$TV_{11}^{OnB_2}$	81.003	129.876	599.000
SN	2.991	3.976	49.000	-	-	-	-

The average user journey length in this data set is rather small, which might be due to it representing the Christmas period. The number and lengths of the user journeys are presented in Table 4.

Table 4. User journey lengths.

Length	1-5	6-10	11-15	16-20	21-25	26-30	31-35	36-40	41-45	46-50
# Users	765980	90872	22595	9305	4663	2363	1296	780	544	398

Analysis Process

The analysis of the user journey data conducted in this study consists of three steps: First, a model for the estimation of non-linearity terms is developed. This step is important, because it allows for a comparison in terms of predictive accuracy between the effects of different model extensions presented in this paper. Second, as it has been observed by Lewis and Reiley (2013), TV effects are modeled as Γ distributed over time to estimate the parameters k and θ of the Γ distribution $\Gamma(k, \theta)$. Finally, the results of the first and second steps are used to estimate the β parameters by applying a hierarchical logistic regression model. These parameters represent the valence and size of the effects resulting from online channel contacts and TV advertising. The following sections describe each step in greater detail. The results are discussed in the next chapter.

Modeling Non-Linearity

The first step of the analysis is to estimate the non-linearity of the covariates from the design matrix (Equation 1). This step is, of course, not a prerequisite to model the impact of TV ads on user journeys. It puts, however, the gain in predictive accuracy for each of the extensions into perspective. Based on the simple multivariate logistic regression model, a Bayesian model is built up using JAGS (Plummer 2003) following Equation 3. The TV covariates are not included in this step to reduce computation times and model complexity.

$$\begin{aligned}
Conv_{ij} &\sim \text{Bernoulli}(Pr_{ij}) \\
Pr_{ij} &= \phi(\hat{X}_{ij}\beta_x) \\
\hat{X}_{ij} &= \{I_0, \dots, I_{PS}, (J_A)^{\gamma_A}, \dots, (J_R)^{\gamma_R}, \dots, (CAS)^{\gamma_{CAS}}\}_{ij} \\
\gamma &\sim \text{Multivariate Normal}(\mu_\gamma, \Omega_\gamma) \\
\beta_x &\sim \text{Multivariate Normal}(\mu_{\beta_x}, \Omega_{\beta_x})
\end{aligned} \tag{3}$$

In this equation, the only difference to the non-hierarchical multivariate logistic regression model is the substitution of X_{ij} with \hat{X}_{ij} , in which each covariate of the design matrix is raised to the $\gamma_{(\cdot)}^h$ power, except for the intercept terms and the time of day variables t, t^2, t^3, t^4 , and WE . The vector γ has the same length as $X_{ij} \setminus \{I_{(\cdot)}, t, t^2, t^3, t^4, WE\}$, denoted as q . The term $\phi(w)$ is the sigmoid function given by $\phi(w) = 1/(1+e^{-w})$, which returns a value between 0 and 1 for a given w . This value is interpreted as the probability Pr for a conversion. In combination with the R package `runjags` (Denwood 2015), the software package JAGS Plummer (2003) is used to obtain the exponents $\gamma_{(\cdot)}$ as well as the β_x parameters. The parameters μ_{β_x} and Ω_{β_x} are the multivariate normal priors for β_x . The μ_{β_x} is a vector of length p , i.e., the length of the row vector X_{ij} , with all entries equal to 0. The prior precision matrix Ω_{β_x} is a $p \times p$ union matrix multiplied by 0.1. The prior μ_γ is a vector of length q with all entries equal to 1. The prior precision matrix Ω_γ is a $q \times q$ union matrix multiplied by 2.5 and, thus, it is acting as a strong prior around 1. In this first analysis, the main interest is the estimation of the γ parameters. They are used in the third step, the hierarchical logistic model, in which the complete set of β parameters is obtained, i.e., $\beta = \{\beta_x, \beta_r\}$.

Modeling TV Effects

In the second step, a model to estimate the time dependency of the TV advertising effect is developed. There are several ways to model this effect. In this paper, two different approaches are presented. First, the TV variable is set to 1 if a spot was broadcast in the last T minutes prior to the online user contact. This approach is easy to implement, but it has also two disadvantages: First, it is very unlikely that the impact of a TV spot in the first few seconds after the broadcast is identical with the impact after 5 minutes, especially if the goal is to explain conversions. Second, the effect instantly collapses after T minutes,

which seems very unrealistic. The second approach proposed here is more realistic, as it takes into account the probability that a customer uses her or his second screen to search for the advertised product grows in the first few minutes after the TV spot has been broadcast. Additionally, it is also considered that the direct impact of a TV spot on the purchasing behavior decreases after a while. This modeling approach is feasible, because, according to previous studies, the uplift of page views after a TV spot has been broadcast looks similar to a Γ distribution (Lewis and Reiley 2013). For this reason, it is proposed to model the TV advertising effect as Γ distributed over time. The Γ distribution can be parameterized by a shape parameter k and a scale parameter θ . The goals of this step is to estimate these values and to obtain the time dependency of the TV effects.

To estimate the parameters, the results from the first modeling step are used to reduce the number of degrees of freedom of the Bayesian model and to reduce computation time. Thus, the β_x values are not estimated in this step, Instead, $z_{ij} = \hat{X}_{ij}\bar{\beta}_x$ is calculated for each contact. The term $\bar{\beta}_x$ represents the median estimate of a simple logistic regression. The Bayesian model is presented in Equation 4.

$$\begin{aligned}
Conv_{ij} &\sim \text{Bernoulli}(Pr_{ij}) \\
Pr_{ij} &= \phi(z_{ij} + \hat{Y}_{ij}\beta_y) \\
\hat{Y}_{ij} &= \{\Gamma(TV_2^{OnB_1}, k, \theta), \dots, \Gamma(TV_{11}^{OnB_2}, k, \theta)\}_{ij} \\
k, \theta &\sim \text{Uniform}(0.01, 10) \\
\beta_y &\sim \text{Multivariate Normal}(\mu_y, \Omega_y)
\end{aligned} \tag{4}$$

In this equation, $\Gamma(TV_s^{(\cdot)}, k, \theta)$ returns the value of the Γ distribution at point $x = TV_s^{(\cdot)}$ with shape parameter k and scale parameter θ . The Γ parameters are sampled from a relatively narrow uniform distribution, because the short term effect of TV advertising is expected to have a maximum in the range between 1 to 60 minutes after the broadcast.

Modeling Cross-Channel Effects

The goal of the third modeling step is to estimate a complete set of β parameters for each contact type. The contact type is represented by the intercept terms $I_{(\cdot)}$ from the design matrix X . The obtained results can be interpreted as cross-channel effects. For instance, the effects of previous contacts with certain channels on conversions that result from search engine requests are represented by β^{os} , i.e., the β parameters for the organic search channel. To achieve this goal, each contact has to be assigned to one of 9 groups representing different contact types. Note that there are only 8 different intercept terms in the design matrix X . The ninth group is related to direct referrals from other websites. A contact of this type is indicated by all other intercept terms being equal to 0, except I_0 .

The modeling approach is driven by several considerations. First, it is known from previous studies that the individual marketing channels influence each other in different ways (Ghose and Yang 2010; Klapdor 2013). Second, it is expected that returning customers behave in a different manner than new customers. For instance, new customers are more likely to click on a search engine advertisement, whereas returning customers, indicated by a higher session number, tend to type-in the URL directly (Rutz et al. 2011). Third, TV ads are expected to have a higher impact on direct type-ins or search requests than on display or affiliate clicks (Naik and Peters 2009).

The hierarchical logistic model with random slopes used here is described by Rossi et al. (2006). The corresponding algorithm `hierRwMNLgit` from the R package `bayesm` involves high computational costs when, as in this case, large sample sizes are used. Therefore, the R package `rpud` by Yau (2015) is used. It contains a parallelized version of the model from the `bayesm` package that uses Graphical Processing Units (GPUs) for computation. This parallelization reduces the computation time by the order of a magnitude. A simplified version without the multinomial part of the original model is presented in Equation 5.

$$\begin{aligned}
Conv_{ij} &\sim \text{Bernoulli}(Pr_{ij}) \\
Pr_{ij} &= \phi(\hat{X}_{ij}\beta_x^G + \hat{Y}_{ij}\beta_y^G) \\
\{\beta_x, \beta_y\}^G &= \Delta + \varepsilon^G
\end{aligned}$$

$$\begin{aligned}\Delta &\sim \text{Multivariate Normal}(\mu_{\Delta}, \Omega_{\Delta}) \\ \varepsilon^G &\sim \text{Multivariate Normal}(\mu_{\varepsilon}, \Omega_{\varepsilon})\end{aligned}\quad (5)$$

In Equation 5, \hat{X}_{ij} and \hat{Y}_{ij} represent the transformed covariates from X_{ij} and Y_{ij} according the first two steps of the analysis. The index $G \in \{A, D, EM, OS, PS, R, SEA, SM, TI\}$ indicates the type of the contact, the vector Δ represents the grand mean of the parameters from all groups, and the residual term ε^G represents the deviation of the individual β^G values from the grand mean Δ .

Results of the Analyses

This chapter is structured as follows: First, the results generated during the first two modeling steps, which were described in the previous chapter, are presented. Second, a comparison of the different modeling approaches is made. It is shown that a model that includes both non-linearity terms as well as TV parameters is the best one to describe the data. Third, the results from the hierarchical logistic regression are presented.

Non-Linearity Parameters

To estimate the γ parameters, the software package JAGS in combination with the R package `runjags` is used to run three MCMC chains with 60,000 burn-in iterations and 125,000 sampling iterations. The sample size is $n=20,000$ in this step. The Gelman-Rubin diagnostic shows convergence of the chains. The results from the non-linearity analysis are presented in Table 5.

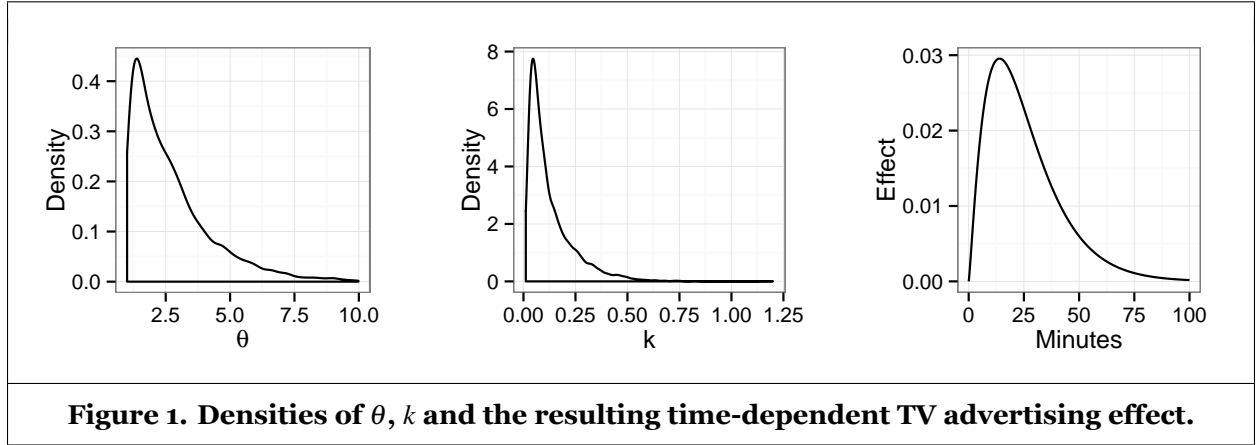
Table 5. Quantiles of the sampled γ parameters.											
Var.	2.5%	25%	50%	75%	97.5%	Var.	2.5%	25%	50%	75%	97.5%
γ_{TI}	0.005	0.058	0.125	0.211	0.401	γ_{KD}	0.042	0.343	0.635	0.984	1.740
γ_A	0.008	0.081	0.162	0.270	0.565	γ_{KSEA}	0.314	0.874	1.097	1.319	1.822
γ_D	0.049	0.336	0.555	0.791	1.410	γ_{KSM}	0.048	0.375	0.675	1.018	1.750
γ_{SEA}	0.003	0.030	0.070	0.128	0.291	γ_{KEM}	0.064	0.407	0.622	0.844	1.305
γ_{SM}	0.022	0.186	0.355	0.590	1.497	γ_{KPS}	0.047	0.393	0.700	1.079	2.004
γ_{EM}	0.006	0.060	0.128	0.221	0.424	γ_{KR}	0.033	0.305	0.553	0.839	1.483
γ_{PS}	0.070	0.455	0.738	1.011	1.574	γ_{SN}	-0.388	0.083	0.144	0.252	0.698
γ_R	0.101	0.490	0.714	0.937	1.477	γ_{IST}	0.001	0.016	0.039	0.078	0.185
γ_{OS}	0.041	0.307	0.566	0.855	1.476	γ_{CWS}	0.091	0.581	0.994	1.443	2.293
γ_{TI}	0.089	0.435	0.617	0.792	1.288	γ_{CAS}	0.504	0.720	0.820	0.913	1.080
γ_{KA}	0.165	0.430	0.574	0.725	1.070	-	-	-	-	-	-

The results clearly show that only counting channel contacts is not an appropriate method for user journey analyses. Most median values are < 1 , which implies saturation effects for each channel. For instance, γ_A , the exponent of the number of previous affiliate marketing contacts during a given session, is close to zero, which means that it makes hardly any difference whether a customer contacted the website once, twice, or ten times through this channel during a session. In this case study, the γ parameter for the additional control variable SN , which was introduced by Chatterjee et al. (2003), shows that this variables has a very small impact since it is always very close to 1.

Time-dependent TV Effect

The densities of the sampled Γ distribution parameters θ and k are presented on the left side and in the middle of Figure 1. They are obtained using 60,000 burn-in steps and 125,000 sampling iterations with $n=10,000$ samples. The graph on the right shows the resulting time-dependent TV spot effect based on

the median values for θ and k . In the minutes after the spot broadcast, the effect strength grows very quickly. It is strongest after approximately 20 minutes and decreases exponentially afterwards. The resulting Γ shaped diagram is in line with the observations from previous TV uplift studies.



Model Comparison

Before the final results are presented, alternative versions of the user journey model discussed above are compared (Table 6). Each model is estimated with and without the TV variables Y and \hat{Y} , respectively. The first two models are based on the approach proposed by Chatterjee et al. (2003) and presented in Table 1. In the first case, the TV variables are set to 1 if the time difference between the spot broadcast and the contact is less than $T = 30$ minutes. In the second case, the Γ parameters θ and k obtained during the second analysis to calculate the time-dependent TV effects are applied. In the third model, the time of day variables and the indicator for weekends WE are added to control for time-dependent conversion rates over the day. If these variables were not included, time-dependent conversion rates might be assigned to the effect of TV spots. In the fourth model, the X variables are transformed into \hat{X} using the γ parameters obtained during the first analysis. Finally, a quadratic term for each covariate is added to account for decreasing awareness effects from individual marketing channels. Table 6 shows the analysis of variance for the ten different models. The residual deviance is given with and without TV variables and the difference of the residual variance to the next smaller model. In addition, the area under the ROC curve (AUC) is reported as a measure for the predictive accuracy of the model. The values are obtained with 2,000,000 contacts for the training set and 1,000,000 contacts for the holdout set.

Table 6. Residual Deviances R for 10 different models and AUC measures.						
Model	R without TV	R with TV	Dev. Delta	Sign.	Dev. Diff.	AUC w. TV
Plain	642409.3	641833.6	575.7	***	-	0.7314
Gamma	642409.3	641612.5	796.8	***	221.1	0.7315
Times	641805.6	641080.1	725.5	***	532.4	0.7326
Non-Linearity	630168.2	629627.7	540.5	***	11452.3	0.7545
Quadratic	625877.2	625368.4	508.8	***	4259.3	0.7624

The results of the model comparison show that the addition of TV variables reduces residual deviance in all cases. Additionally, they show that the time-dependent TV effects reduce more deviance than the simpler approach to model TV effects. The addition of variables to describe the time of the day also reduces deviance. However, most of the deviance is reduced when the γ parameters are applied to transform X into \hat{X} . Even the addition of a quadratic term for each covariate has a much higher impact on residual deviance than the TV effects. This shows the importance of the feature selection process in user journey modeling, which is, however, not the focus of this study.

Results from the Non-Hierarchical Logistic Regression

The results from the hierarchical model cannot be presented in great detail here, because one set of β values for each contact type was obtained. However, to offer an overview of the size of the effects, the complete set of β parameters for the non-hierarchical logistic model in Table 7 are presented. The used model is the one with included TV effects and non-linearity from the fourth row in Table 6. Note that the \hat{X} and \hat{Y} covariates were standardized to allow for easy comparison between individual effect sizes. Thus, the values listed in Table 7 show the increase of the logit of Pr by $\beta_{(\cdot)}$ resulting from adding one standard deviation of the respective covariate.

Table 7. Results from the simple logistic regression with transformed covariates.									
$\beta_{(\cdot)}$	Est.	<i>SD</i>	<i>z</i>	Sig. Lv.	$\beta_{(\cdot)}$	Est.	<i>SD</i>	<i>z</i>	Sig. Lv.
0	-3.461	0.004	-769.985	***	CWS	0.006	0.002	2.487	*
I_{OS}	0.182	0.012	15.367	***	CAS	0.218	0.003	72.476	***
I_{TI}	0.413	0.013	31.175	***	t	1.266	0.090	14.004	***
I_A	0.489	0.009	54.591	***	t^2	-4.812	0.387	-12.441	***
I_D	-0.016	0.008	-1.998	*	t^3	6.262	0.548	11.424	***
I_{SEA}	0.308	0.015	19.906	***	t^4	-2.783	0.252	-11.061	***
I_{SM}	0.017	0.005	3.088	**	WE	-0.030	0.004	-7.482	***
I_{EM}	0.273	0.009	29.778	***	$TV_2^{OnB_1}$	0.008	0.004	2.084	*
I_{PS}	0.039	0.005	7.291	***	$TV_4^{OnB_1}$	0.017	0.004	4.464	***
J_{TI}	0.195	0.002	80.345	***	$TV_6^{OnB_1}$	0.010	0.004	2.458	*
J_A	0.122	0.003	41.561	***	$TV_{11}^{OnB_1}$	0.027	0.004	6.211	***
J_D	0.061	0.003	17.897	***	$TV_5^{OffP_1}$	0.008	0.004	2.089	*
J_{SEA}	0.223	0.004	56.486	***	$TV_7^{OffP_1}$	0.024	0.004	5.773	***
J_{SM}	0.006	0.004	1.542		$TV_8^{OffP_1}$	-0.008	0.004	-1.875	.
J_{EM}	0.146	0.003	51.579	***	$TV_1^{OffB_1}$	-0.002	0.004	-0.617	
J_{PS}	0.029	0.003	9.241	***	$TV_2^{OffB_1}$	-0.009	0.004	-2.095	*
J_R	0.055	0.003	16.093	***	$TV_3^{OffB_1}$	-0.016	0.004	-4.044	***
K_{OS}	-0.056	0.005	-11.882	***	$TV_4^{OffB_1}$	-0.013	0.004	-3.259	**
K_{TI}	-0.118	0.009	-13.263	***	$TV_5^{OffB_1}$	-0.049	0.004	-11.469	***
K_A	-0.168	0.004	-40.109	***	$TV_6^{OffB_1}$	0.035	0.004	9.697	***
K_D	0.006	0.004	1.466		$TV_9^{OffB_1}$	-0.001	0.004	-0.302	
K_{SEA}	-0.078	0.006	-12.301	***	$TV_{10}^{OffB_1}$	0.009	0.004	2.312	*
K_{SM}	-0.015	0.005	-3.060	**	$TV_{11}^{OffB_1}$	0.002	0.004	0.496	
K_{EM}	-0.028	0.005	-6.058	***	$TV_1^{OnB_2}$	0.015	0.004	4.037	***
K_{PS}	-0.033	0.004	-7.619	***	$TV_2^{OnB_2}$	0.011	0.004	2.809	**
K_R	-0.053	0.005	-10.299	***	$TV_3^{OnB_2}$	0.009	0.004	2.220	*
SN	-0.334	0.014	-24.062	***	$TV_5^{OnB_2}$	0.040	0.004	10.317	***
IST	0.728	0.006	116.306	***	$TV_{11}^{OnB_2}$	0.014	0.004	3.769	***

The intercept β_0 and the term β_{SN} indicate that the baseline conversion probability is, at approximately $p \approx \phi(-3.461 - SN^{TSN} \cdot 0.334) \approx 0.02$, very low. As stated before, this value corresponds to conversion

rates for direct referrals. This baseline probability for conversions matches typical conversion rates found in the industry and is in line with previous studies (Ghose and Yang 2010). The probabilities for conversions through affiliate channel, search engine channel, e-mail channel, and direct type-ins are higher than this baseline probability, which is represented by the positive value of the respective $\beta_{I(\cdot)}$ terms.

The effect of the number of previous channel contacts during one session $\beta_{I(\cdot)}$ is positive in all cases, although the value for contacts through the social media channel is not significant. This indicates that customers who are more likely to convert tend to use multiple marketing channels during a given session before they make their purchase decision.

The effect of the number of channel contacts in previous sessions $\beta_{K(\cdot)}$, however, is negative for all channels except for the display channel effect β_{K_D} , which is not significant. At first, this finding is unexpected, because it is generally assumed that returning customers with more than one session are more likely to convert than customers with shorter user journeys. As the data set has been generated during the Christmas season, one could argue that customers shopping during the season tend to use the shop for spontaneous gift purchases that can be completed during one user session.

The effects of TV spots that advertise for the online shop (i.e., OnB_1 and OnB_2) are positive in all cases. In addition, the offline and product-related spots have a positive impact on the conversion probability, except $TV_8^{OffP_1}$ which does not fulfill the significance criteria.

However, the effect from the TV spots is rather low in comparison to the effect from the online channel contacts. This is because it is not known whether a user in fact watched a certain TV program. Additionally, information about the reach and the placements of the spots within the advertising blocks is unavailable. Consequently, to develop the model it is assumed that the TV spots have an effect on each contact of every customer. Thus, the β_{TV} values represent the mean TV effect for all customers.

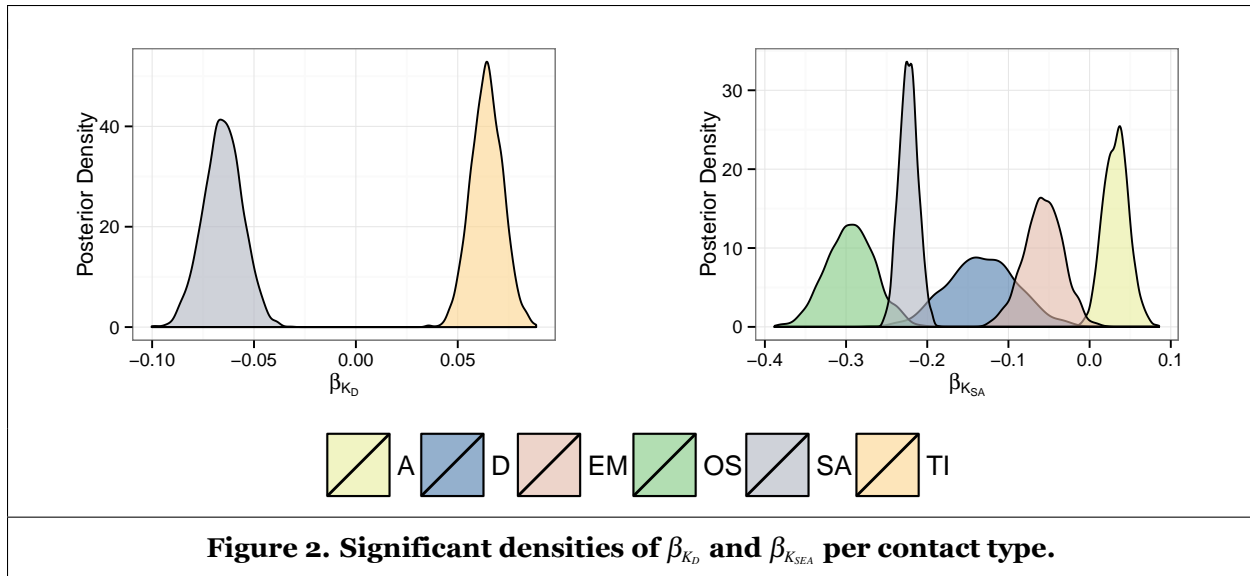
Depending on the TV station, there are positive or negative effects resulting from offline brand spots, i.e., $OffB_1$. Significant negative effects from TV spots are rather unintuitive, because one would expect they either have a positive effect, if the customer watched the spot, or no effect at all, if the customer did not watch it. However, negative effects can be explained with customers who make last-minute decisions. They tend to search online for a certain product to read reviews, for instance. However, these customers tend to buy the product in a retail store, because they might worry that the product cannot be delivered in time. This hypothesis is realistic, because a multitude of the brand-related offline spots have been broadcast in the few days before Christmas Eve. Furthermore, there might be lots of customers who visit the online shop to read reviews about a certain product and, nevertheless, decide to visit the retail store to try out the product before purchasing it (Cheema and Papatla 2010).

Cross-Channel Effects

Running the hierarchical logistic model from the R package `rpud` with 2,000,000 samples and 250,000 iterations results in $9 \cdot 48 = 432$ posterior densities of the $\beta_{(\cdot)}^c$ parameters. These values cannot be reported in full detail here. Instead, the results are discussed by example. Some significant effects of the covariates from $X \setminus I(\cdot)$ on the contact types represented by $I(\cdot)$ are shown in Figure 2 and 3.

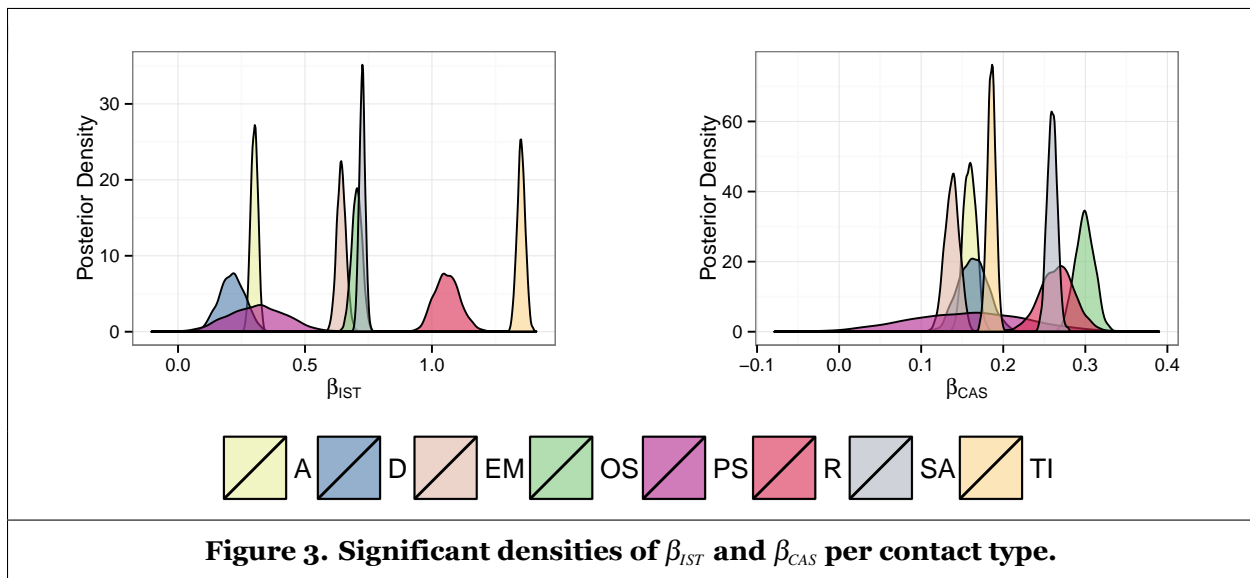
The graph included on the left side of Figure 2 shows the effect of clicks on display ads in previous sessions on search engine advertising contacts and direct type-in contacts. The non-hierarchical logistic regression did not reveal significant results for β_{K_D} . However, the hierarchical analysis shows that $\beta_{K_D}^{SEA}$ and $\beta_{K_D}^{TI}$ are significant. The result is relatively intuitive, because previous contacts with display ads increase the awareness for the brand and, hence, increase the probability for direct type-ins (Chatterjee 2008). The probability for conversions after clicks on search engine ads decreases with increasing K_D . The negative value for $\beta_{K_D}^{SEA}$ is sensible, since one of the major goals of search engine advertising is the acquisition of new customers (Rutz et al. 2011).

The graph included on the right side of Figure 2 shows the effect of the number of search engine advertising contacts in previous sessions K_{SEA} on five different channels. It has a positive impact on the probability for conversions through the affiliate channel and negative impacts on the probability for conversions through the organic search, the search engine, the display, and the e-mail marketing channels. A negative value for $\beta_{K_{SEA}}$ is in line with the results from the non-hierarchical analysis. However, the hier-

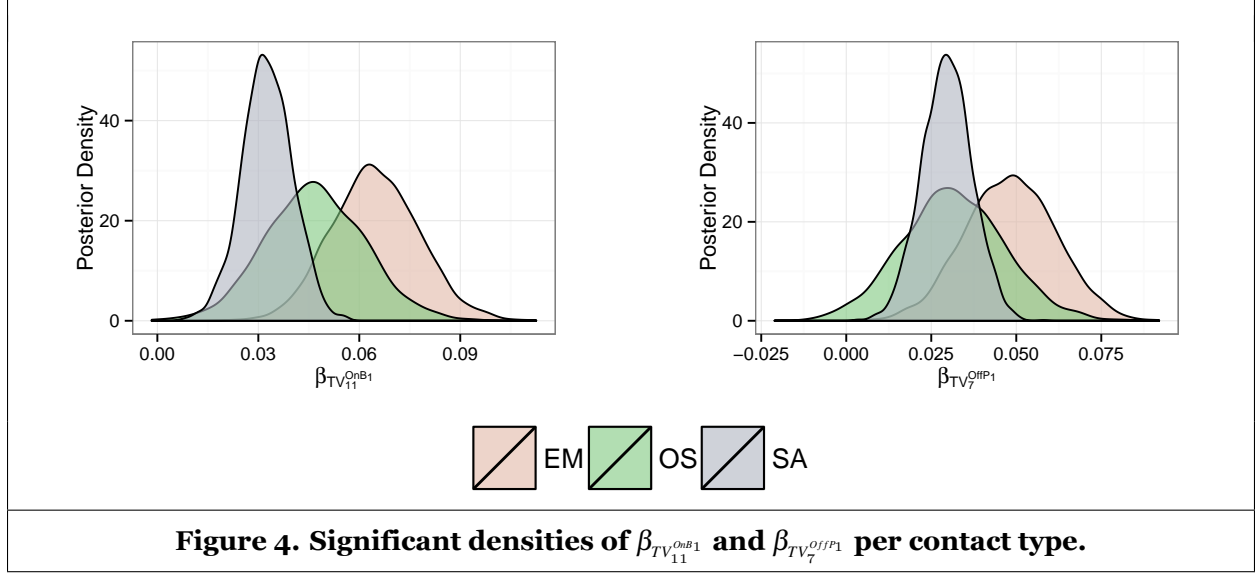


archical analysis reveals different effects across contact types. The affiliate channel, for instance, benefits from the increasing value for K_{SEA} . One reason for this might be the growing awareness of the customers for discount coupons that are often offered in affiliate marketing.

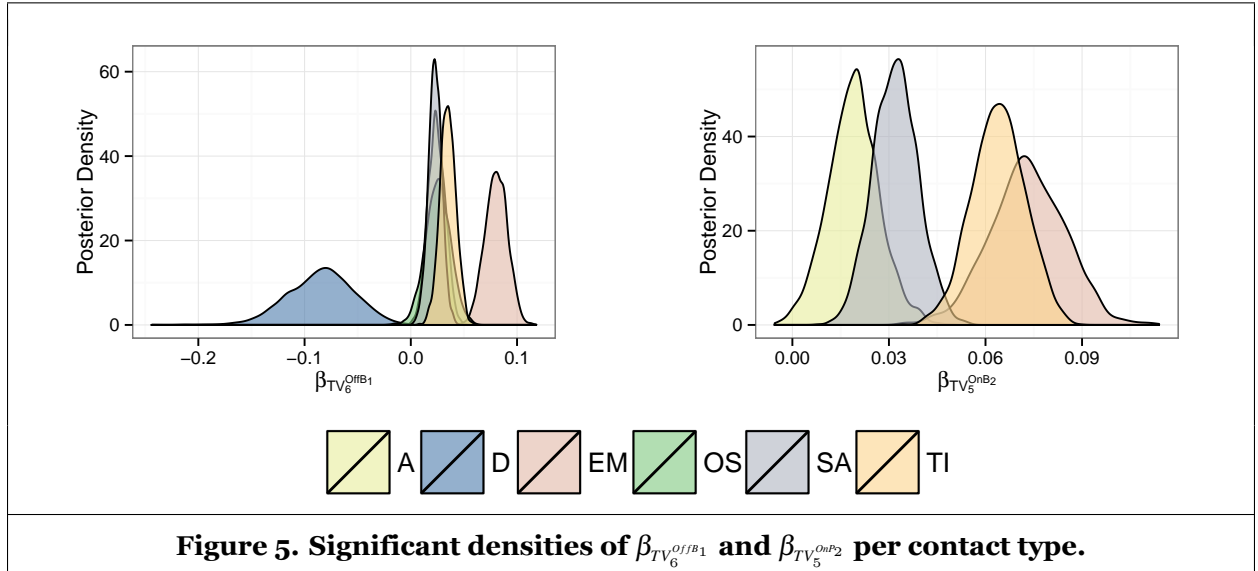
The graph included on the left side of Figure 3 shows the effect of the inter-session time IST on eight different channels. The hierarchical analysis shows that the deviation from the mean effect $\beta_{IST} = 0.728$, calculated in the non-hierarchical logistic regression, is relatively high. The inter-session time has a high impact on the probability for conversions after a direct type-in and a low impact for a conversion after a display click. The effect of the number of conversions in previous sessions CAS on eight different channels is shown in on the right side of Figure 3. The CAS value has the highest impact on the probability for conversions after an organic search request and the less impact for a conversion after an e-mail advertising click.



Significant effects of TV spots are shown in Figure 4 and 5 by example. Non-significant values are omitted for clarity. The graph included on the left side of Figure 4 shows the impact of the online brand spot #1, station #11, on the probability for e-mail conversions, organic search conversions and search engine advertising conversions. The increase of the probability for search engine conversions after a TV spot is



relatively intuitive and in line with previous studies (e.g., Kitts et al. 2010; Zigmond and Stipp 2010). The interpretation of the value for $\beta_{TV_{11}^{OnB1}}^{EM}$ needs more explanation, because the increase of the probability for e-mail conversions after a TV spot is rather surprising. The reason might be the concurrent e-mail marketing campaign that was dominated by providing discounts coupons for online purchases. Apparently, customers who had received such a coupon have been likely to reopen that e-mail and use the coupon to purchase a product after watching a brand-related TV spot. This finding is consistent for all TV spots and has managerial implications that are discussed in the next chapter. The effect of an offline product-related spot is presented on the right side of Figure 4 and shows a similar picture as the online-related brand spot #1. The effect of direct type-ins is not significant in this case because it is rather uncommon to directly type-in the product-related URL in the browser.



The left side of Figure 5 shows the cross-channel effects for the offline brand spot on station #6. As seen before, the highest positive impact of the TV spot can be observed for e-mail conversions, followed by the impact on organic search conversions and search engine advertising conversions. Additional significant positive effects are observed for affiliate conversions and direct type-ins. The increase of the probability of affiliate conversions might be caused by the tendency of some customers to search for discount vouchers

before visiting the website and purchasing a product. The graph included on the right side of Figure 5 presents the effect of the second online brand spot on station #5. It draws a very similar picture. The only difference is the missing negative effect on display conversions. This might be again due to the fact that the advertiser also operates retail stores and cross-channel saturation effects (Piercy 2012).

In summary, the application of the hierarchical model provides new insights into cross-channel effects that determine the user behavior. The AUC for this model is 0.7606 and thereby higher than the analog model without the quadratic terms from the non-hierarchical analysis. In particular in the context of real-time advertising, where millions of decisions are made every second, such a model is valuable to determine the size of the bids on the basis of the predicted probability for a conversion.

Implications, Limitations and Outlook

In this paper, a model commonly used for clickstream analyses is extended by introducing offline advertising effects. Such an extended model outperforms the simple clickstream model in terms of predictive accuracy. Adding a hierarchical structure, the cross-channel effects of both online and TV advertising activities are revealed.

Implications

The approach presented in this study has several implications for researchers in the fields IS and marketing. First, future research needs to be done on decision support systems in e-commerce that integrate tracking data and offline advertising data from other data sources to improve the quality of strategic decisions, for example the allocation of parts of the marketing budget. Since customer behavior is influenced by many external factors such as seasonal effects (as in this case study) or new brands or products, analyses based on user data should be updated frequently to obtain reliable results. Therefore, efficient tools to manage and analyze these high-volume tracking data and data from offline sources are required to make the model viable in practice, for example in strategic decision support systems or bidding agents in real-time advertising (Stange and Funk 2014). Second, for future research at the intersection of IS and marketing, data from tracking systems that identify users across different devices should be used to arrive at a better understanding of users' behavior in e-commerce contexts. In combination with data from traditional offline advertising channels such as print and television advertising, the model presented in this paper can provide even more precise insights into cross-device and cross-media user behavior. Third, marketing researchers can use this approach to analyze the impact of additional properties of offline advertising, such as information about television stations, the positions of the spots within the advertising blocks, or the context in which the spot was aired. These model extensions can yield better predictive accuracies and improve strategic decision support, for example for future advertising campaigns.

In addition, the study has several managerial implications: First, a model to estimate time dependency of the short-term TV spot effects is proposed. The results show that the strongest effect is observed 15 minutes after a spot has been broadcast. Bidding agents in real-time advertising can use this information to select ads related to the product advertised within that time frame to optimize profits from cross-channel marketing. Second, the model proposed here can be used in practice to plan future online advertising campaigns, i.e., the allocation of future marketing budgets. For instance, advertising activities on channels that show a relatively high conversion rate (in the case study, affiliate marketing, search engine advertising) could be expanded. The same is valid for TV advertising activities: The spendings for TV ads on stations that show a high positive effect on the purchase probability could be increased whereas the number of spots on stations without a noticeable effect might be reduced. To optimize the budget the estimated effects per channel have to be weighted in light of the costs per acquisition for each channel. Third, the results of the hierarchical analysis show that the impact of TV advertisement is different for each online channel. Companies who consider a TV advertising campaign should coordinate advertising activities on online channels that greatly benefit from TV advertising activities, for example search engine advertising, organic search, and e-mail advertising. For instance, if an advertiser runs a campaign for a certain product on television, it would make sense to run a search engine advertising campaign during that time to increase the spillover effect from the TV ad. Surprisingly, in this study, the highest impact of TV ads on the conversion probability can be observed for the e-mail channel. During an advertising campaign for a certain product on TV, an advertiser should also conduct a complementary e-mail campaign

to optimize the spillover effect from the TV channel. This is but one example emphasizing the importance of a well-balanced cross-channel marketing strategy.

Limitations

Although it clearly reveals the expected TV effects, the analysis has primarily seven limitations. These are mostly related to incomplete data. First, to model the impact of television ads, it is crucial to have an idea of how many customers might have been watching a certain TV station at a given time. While it is possible to determine when a spot was aired and where individual spots were placed in the advertising blocks, it is hard to acquire reliable data on how many people actually saw these ads. This information should be still included in future analyses. Second, the device information of each contact was unavailable. However, this information is important, because it is expected that the probability for a conversion on a mobile tablet, which is used as a second device while watching TV, is higher than for a conversion on a standard PC, which is usually not instantly available when watching TV. Future studies in this area need to include information about user devices to build a model that can validate this hypothesis. Third, to estimate the TV impact, a simple logistic regression and a hierarchical logistic regression model were used. These models do not restrict the impact of TV spots in any form, which might be useful in some cases. For instance, it is rather unlikely that the impact of TV ads is less than zero. Therefore, it could make sense to change the prior information of the TV effects β_{TV} from multivariate normal to truncated multivariate normal. However, efficient algorithms to estimate the parameters for this kind of model are currently unavailable. Fourth, the focus was on conversions here and the activities of a user on the retailer's website were ignored. However, it would also be very important to analyze onsite user behavior with a focus on products and purchase intentions depending on TV advertising activities. Fifth, the increase in the AUC value resulting from the addition of the parameters for TV spots and from the hierarchical model structure were small in comparison to the increase that resulted from the introduction of the γ exponents and the quadratic terms. This fact emphasizes that the feature selection process is more important than a sophisticated model structure. Sixth, the data was collected during Christmas time and, thus, includes many shorter user journeys suggesting the spontaneous purchase of gifts. Although the modeling approach presented here is generally applicable for offline-online studies, the results from the case study can not be transferred to other periods without limitation. Seventh, the focus of this study was on the short-term effects of TV advertising activities. As indicated by previous studies, TV advertising focuses, however, on the awareness of potential customers (Liaukonyte et al. 2015). It was not possible to model the awareness effect, because user journeys without the influence of a TV campaign are not contained in the data set. For experimental settings in future studies, it would be desirable to measure both short-term and long-term TV advertising effects.

Outlook

The analysis of cross-channel marketing effects is often limited to online marketing channels, because they enable advertisers to track every customer activity. However, online marketing is only one part of the broad range of marketing activities that, ideally, are seamlessly coordinated. The analysis of the effect of online-offline marketing activities is a chance for practitioners to improve their marketing activities. For instance, an important question for practitioners is how to allocate parts of the budget to individual channels. The estimations of the proposed model can be used to answer this question. In addition, the measurement of online-offline advertising effects offer many opportunities for future scholarship at the intersection of IS and marketing. This analysis represents a first step in this direction.

References

- Anderl, E., Becker, I., Wangenheim, F., and Schumann, J. H. 2014. "Mapping the Customer Journey: A Graph-Based Framework for Online Attribution Modeling," *Available at SSRN: <http://ssrn.com/abstract=2343077>*.
- Bolten, B. 2014. "Screen Life 2014," Retrieved from http://www.ipsos.de/assets/files/presse/2014/publikationen/Ipsos_Screen%20Life%202014_Praesentation_final.pdf, January 17, 2015.
- Bucklin, R., and Sismeiro, C. 2009. "Click here for Internet insight: Advances in clickstream data analysis in marketing," *Journal of Interactive Marketing* (23:1), pp. 35–48.
- Bucklin, R. E., and Sismeiro, C. 2003. "A Model of Web Site Browsing Behavior Estimated on Clickstream Data," *Journal of Marketing Research* (40:3), pp. 249–267.
- Chatterjee, P. 2008. "Are unclicked ads wasted? Enduring effects of banner and pop-up ad exposures on brand memory and attitudes," *Journal of electronic commerce Research* (9:1), pp. 51–61.
- Chatterjee, P. 2010. "Multiple-channel and cross-channel shopping behavior: role of consumer shopping orientations," *Marketing Intelligence & Planning* (28), pp. 9–24.
- Chatterjee, P., Hoffman, D., and Novak, T. 2003. "Modeling the clickstream: Implications for web-based advertising efforts," *Marketing Science* (22:4), pp. 520–541.
- Cheema, A., and Papatla, P. 2010. "Relative importance of online versus offline information for Internet purchases: Product category and Internet experience effects," *Journal of Business Research* (63:9), pp. 979–985.
- Cho, C.-H. 2003. "Factors influencing clicking of banner ads on the WWW," *Cyberpsychology & Behavior* (6:2), pp. 201–215.
- Courtois, C., and D'heer, E. 2012. "Second screen applications and tablet users: constellation, awareness, experience, and interest," *Proceedings of the 10th European Conference on Interactive TV and Video*, pp. 153–156.
- Dalessandro, B., and Perlich, C. 2012. "Causally motivated attribution for online advertising," *Proceedings of the 6th International Workshop on Data Mining for Online Advertising and Internet Economy*.
- Denwood, M. J. 2015. "R Package runjags V2.0.1-2," Retrieved from <http://cran.r-project.org/web/packages/runjags/index.html>, April 18, 2015.
- Dinner, I., van Heerde, H. J., and Neslin, S. A. 2014. "Driving Online and Offline Sales: The Cross-channel Effects of Traditional, Online Display, and Paid Search Advertising," *Journal of Marketing Research* (51:5), pp. 527–545.
- Ghose, A., and Yang, S. 2010. "Analyzing the Relationship Between Organic and Sponsored Search Advertising: Positive, Negative, or Zero Interdependence?," *Marketing Science* (29:4), pp. 602–623.
- Joo, M., Wilbur, K. C., and Zhu, Y. 2014. "Effects of TV Advertising on Keyword Search," *Management Science* (60:1), pp. 56–73.
- Jordan, P., Mahdian, M., Vassilvitskii, S., and Vee, E. 2011. "The multiple attribution problem in pay-per-conversion advertising," in: *Algorithmic Game Theory*, G. Persiano (ed.), Berlin: Springer, pp. 31–43.
- Kitts, B., Wei, L., Au, D., Powter, A., and Burdick, B. 2010. "Attribution of conversion events to multi-channel media," *Proceedings of the IEEE International Conference on Data Mining*, pp. 881–886.
- Klapdor, S. 2013. *Effectiveness of Online Marketing Campaigns*, Wiesbaden: Springer.
- Lee, K., Orten, B., Dasdan, A., and Li, W. 2012. "Estimating conversion rate in display advertising from past performance data," *Proceedings of the 18th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp. 768–776.

- Lewis, R. A., and Reiley, D. H. 2013. "Down-to-the-Minute Effects of Super Bowl Advertising on Online," *14th ACM Conference on Electronic Commerce* (9:4), pp. 1–22.
- Liaukonyte, J., Teixeira, T., and Wilbur, K. C. 2015. "Television Advertising and Online Shopping," *Marketing Science* (34:3), pp. 311–330.
- Lodish, L., Abraham, M., Livelsberger, J., Lubetkin, B., Richardson, B., and Stevens, M. E. 1995. "A Summary of Fifty-Five In-Market Experimental Estimates of the Long-Term Effect of TV Advertising," *Marketing Science* (14:3), pp. 133–140.
- Naik, P., and Peters, K. 2009. "A Hierarchical Marketing Communications Model of Online and Offline Media Synergies," *Journal of Interactive Marketing* (23:4), pp. 288–299.
- Nottorf, F. 2014. "Modeling the clickstream across multiple online advertising channels using a binary logit with bayesian mixture of normals," *Electronic Commerce Research and Applications* (13:1), pp. 45–55.
- Olbrich, R., and Schultz, C. D. 2014. "Multichannel advertising: does print advertising affect search engine advertising?" *European Journal of Marketing* (48:9/10), pp. 1731–1756.
- Piercy, N. 2012. "Positive and negative cross-channel shopping behaviour," *Marketing Intelligence & Planning* (30:1), pp. 83–104.
- Plummer, M. 2003. "JAGS: A program for analysis of Bayesian graphical models using Gibbs sampling," *Proceedings of the 3rd International Workshop on Distributed Statistical Computing*, pp. 125–133.
- Richardson, M., Dominowska, E., and Ragno, R. 2007. "Predicting clicks: estimating the click-through rate for new ads," *Proceedings of the 16th International Conference on World Wide Web*, pp. 521–530.
- Rossi, P., Allenby, G., and McCulloch, R. 2006. *Bayesian statistics and marketing*, Wiley.
- Rutz, O. J., Trusov, M., and Bucklin, R. E. 2011. "Modeling Indirect Effects of Paid Search Advertising: Which Keywords Lead to More Future Visits?" *Marketing Science* (30:4), pp. 646–665.
- Shao, X., and Li, L. 2011. "Data-driven multi-touch attribution models," *Proceedings of the 17th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp. 258–264.
- Stange, M., and Funk, B. 2014. "Real-Time-Advertising," *Business & Information Systems Engineering* (56:5), pp. 305–308.
- Stange, M., and Funk, B. 2015. "How Much Tracking Is Necessary - The Learning Curve in Bayesian User Journey Analysis," *Proceedings of the 23rd European Conference on Information Systems*.
- Veit, D., Clemons, E., Benlian, A., Buxmann, P., Hess, T., Kundisch, D., Leimeister, J. M., Loos, P., and Spann, M. 2014. "Business Models," *Business & Information Systems Engineering* (6:1), pp. 45–53.
- Yau, C. 2015. "R Package RPUD 0.5.0," Retrieved from <http://www.r-tutor.com/gpu-computing>, January 07, 2015.
- Zigmond, D., and Stipp, H. 2010. "Assessing a New Advertising Effect Measurement of the Impact of Television Commercials on Internet Search Queries," *Journal of Advertising Research* (50), pp. 162–168.